https://doi.org/10.5194/gmd-2020-55
Preprint. Discussion started: 17 April 2020

Geoscientific Model Development Discussions

Detection of Atmospheric Rivers with Inline Uncertainty Quantification: TECA-BARD v1.0

Travis A. O'Brien[1,2], Mark D. Risser[2], Burlen Loring[3], Abdelrahman A. Elbashandy[3], Harinarayan Krishnan[3], Jeffrey Johnson[4], Christina M. Patricola[2], John P. O'Brien[2,5], Ankur Mahesh[2], Prabhat[6], Sarahí Arriaga Ramirez[2,7], Alan M. Rhoades[2], Alexander Charn[2,8], Héctor Inda Díaz[2,7], and William D. Collins[2,8]

[1]Dept. of Earth and Atmospheric Sciences, Indiana University, Bloomington, Indiana, USA
[2]Climate and Ecosystem Sciences Division, Lawrence Berkeley National Lab, Berkeley, California, USA
[3]Computational Sciences Division, Lawrence Berkeley National Lab, Berkeley, California, USA
[4]Cohere Consulting LLC, Seattle, WA, USA
[5]Dept. of Earth and Planetary Science, University of California, Santa Cruz, California, USA
[6]National Energy Research Scientific Computing Center, Lawrence Berkeley National Lab, Berkeley, California, USA
[7]Dept. of Land, Air and Water Resources, University of California, Davis, California, USA
[8]Dept. of Earth and Planetary Science, University of California, Berkeley, California, USA

Correspondence: Travis A. O'Brien (obrienta@iu.edu)

Abstract. It has become increasingly common for researchers to utilize methods that identify weather features in climate models. There is an increasing recognition that the uncertainty associated with choice of detection method may affect our scientific understanding. For example, results from the Atmospheric River Tracking Method Intercomparison Project (ARTMIP) indicate that there are a broad range of plausible atmospheric river (AR) detectors, and that scientific results can depend on the algorithm used. There are similar examples from the literature on extratropical cyclones and tropical cyclones. It is therefore imperative to develop detection techniques that explicitly quantify the uncertainty associated with the detection of events. We seek to answer the question: given a 'plausible' AR detector, how does uncertainty in the detector quantitatively impact scientific results? We develop a large dataset of global AR counts, manually identified by a set of 8 researchers with expertise in atmospheric science, which we use to constrain parameters in a novel AR detection method. We use a Bayesian framework to sample from the set of AR detector parameters that yield AR counts similar to the expert database of AR counts; this yields a set of 'plausible' AR detectors from which we can assess quantitative uncertainty. This probabilistic AR detector has been implemented in the Toolkit for Extreme Climate Analysis (TECA), which allows for efficient processing of petabyte-scale datasets. We apply the TECA Bayesian AR Detector, TECA-BARD v1.0, to the MERRA2 reanalysis and show that the sign of the correlation between global AR count and El Niño Southern Oscillation depends on the set of parameters used.

1 Introduction

There is a growing body of literature in which researchers decompose precipitation and other meteorological processes into constituent weather phenomena, such as tropical cyclones, extratropical cyclones, fronts, mesoscale convective systems, and atmospheric rivers (e.g., Kunkel et al., 2012; Neu et al., 2013; Walsh et al., 2015; Schemm et al., 2018; Zarzycki et al., 2017;

https://doi.org/10.5194/gmd-2020-55
Preprint. Discussion started: 17 April 2020


Geoscientific Model Development Discussions

Wehner et al., 2018). Research focused on atmospheric rivers (ARs) in particular has contributed a great deal to our understanding of the water cycle (Zhu and Newell, 1998; Sellars et al., 2017), atmospheric dynamics (Hu et al., 2017), precipitation variability (Dong et al., 2018), precipitation extremes (Dong et al., 2018), impacts, meteorological controls on the cryosphere (Gorodetskaya et al., 2014; Huning et al., 2017, 2019), and uncertainty in projections of precipitation in future climate change scenarios (Gershunov et al., 2019).

Over the past decade, there has been a growth in the number of methods used to detect atmospheric rivers, and in the last five years there has been a growing recognition that uncertainty in AR detection may impact our scientific understanding; the Atmospheric River Tracking Method Intercomparison Project (ARTMIP) was created to assess this impact (Shields et al., 2018). Through a series of controlled, collaborative experiments, results from ARTMIP have shown that at least some aspects of our understanding of AR-related science indeed depend on detector design (Shields et al., 2018; Rutz et al., 2019). Efforts related to ARTMIP have similarly shown that some aspects of AR-related science depend on the detection algorithm used (Huning et al., 2017; Ralph et al., 2019).

ARTMIP has put significant effort into quantifying uncertainty, and the community is poised to imminently produce several important papers on this topic. It would be impractical to perform ARTMIP-like experiments for every AR-related science question that arises, which raises the question of how best to practically deal with uncertainty in AR detection.

This uncertainty arises because there is no theoretical and quantitative definition of an AR. Only recently did the community come to a consensus on a qualitative definition (Ralph et al., 2018). In order to do quantitative science related to ARs, researchers have had to independently form quantitative methods to define atmospheric rivers (Shields et al., 2018). Existing AR detection algorithms in the literature are predominantly heuristic: e.g., they consist of a set of rules used to isolate ARs in meteorological fields. Inevitably, heuristic algorithms also contain unconstrained parameters (e.g., thresholds). Across phenomenon-detection literature (ARs and other phenomena), the prevailing practice is for a researcher to use expert judgement to select these parameters. The only exception of which the authors are aware is that of Zarzycki and Ullrich (2017), who use an optimization method to determine parameters for a tropical cyclone (TC) detector.

Even if one were to adopt a similar optimization framework for an AR detector, this still would not address the issue that uncertainty in AR detection can qualitatively affect scientific results. This sort of problem has motivated the use of formal uncertainty quantification frameworks, in which an ensemble of 'plausible' AR detectors are run simultaneously. However, these frameworks need data against which to assess the plausibility of a given AR detector. Zarzycki and Ullrich (2017) were able to take advantage of an existing, human-curated dataset of TC tracks. No such dataset exists for ARs.

A key challenge for developing such a dataset is the human effort required to develop it. The best type of dataset would be one in which experts outline the spatial footprints of ARs; this would be tedious and time consuming to produce. A simpler alternative would be to simply identify the number of ARs in a set of given meteorological fields. Even though a dataset of AR counts is perhaps less helpful than a dataset of AR footprints, we hypothesize that such a dataset could serve to constrain the parameters in a given AR detector.

This manuscript addresses both of these challenges: we develop a formal Bayesian framework for sampling 'plausible' sets of parameters from an AR detector, and we develop a database of AR counts with which to constrain the Bayesian method.

https://doi.org/10.5194/gmd-2020-55
Preprint. Discussion started: 17 April 2020

Geoscientific Model Development Discussions

We provide a general outline for the Bayesian framework as well as a specific implementation: the Toolkit for Extreme Climate Analysis Bayesian AR Detector version 1.0 (TECA-BARD v1.0; Section 2). We show that TECA-BARD v1.0 performs comparably to an ensemble of algorithms from ARTMIP and that it emulates the counting statistics of the contributors who provided AR counts (Section 3). We demonstrate that answers to the question *Are there more ARs during El Niño events?* depends qualitatively on the set of detection parameters (Section 4).

2 The Bayesian approach

2.1 Overview

We start with a general description of how a Bayesian framework, in combination with a dataset of AR counts, can be applied to an AR detector. We consider a generic heuristic detection algorithm with tunable parameters $\boldsymbol{\theta}$ (thresholds, etc.) that, when given an input field $\boldsymbol{Q}$ (e.g., integrated vapor transport, IVT), can produce a count of ARs within that field. For compactness, we will represent this heuristic algorithm and subsequent counting as a function $f(\boldsymbol{\theta}|\boldsymbol{Q})$. That is, for a given field $\boldsymbol{Q}$ and a specific choice of tuning parameters, f returns the number of detected ARs in $\boldsymbol{Q}$.

Further, we assume that we have an dataset of M actual AR counts, denoted by N, associated with a set of independent input fields (i.e., generated by an expert counting the ARs; see Section 2.2): $\{(N_i, \boldsymbol{Q}_i) : i = 1, \ldots, M\}$. With a quantitatively defined prior on the tunable parameters $p_\theta(\boldsymbol{\theta})$, we can use Bayes' theorem to define the posterior probability of $\boldsymbol{\theta}$ given the AR counts $\boldsymbol{N}$ and input fields $\boldsymbol{Q}$:

$$p(\boldsymbol{\theta} \mid \boldsymbol{N}, \boldsymbol{Q}) \propto \frac{\left(\prod_{i=1}^{M} \mathcal{L}(N_i \mid \boldsymbol{\theta}, \boldsymbol{Q}_i) \right) \cdot p_\theta(\boldsymbol{\theta})}{p_N(\boldsymbol{N})} \tag{1}$$

We propose to base the likelihood $\mathcal{L}$ on counts from the heuristic model $N_i' = f(\boldsymbol{\theta}, \boldsymbol{Q}_i)$. We model $\mathcal{L}$ as a normal distribution centered on N_i':

$$\mathcal{L}(N_i \mid \boldsymbol{\theta}, \boldsymbol{Q}_i) = \mathcal{N}(N_i \mid N_i', \sigma), \tag{2}$$

where σ is a nuisance parameter that is ultimately integrated over. While the normal distribution is typically assigned to a continuous (real-valued) variable, here we simply use it as a quantitative way to minimize the squared error between each N_i and N_i'.

2.2 A Database of Expert AR Counts

In order to constrain a Bayesian AR detection algorithm, we developed a database of global AR counts. We designed a simple graphical user interface (GUI) that displays a meteorological plot, as shown in Figure 1, for a given instance of time. Times are chosen randomly within the years 2008 and 2009. The interface allows a user to enumerate ARs within a given field by clicking

https://doi.org/10.5194/gmd-2020-55
Preprint. Discussion started: 17 April 2020

Geoscientific Model Development Discussions
Open Access
EGU

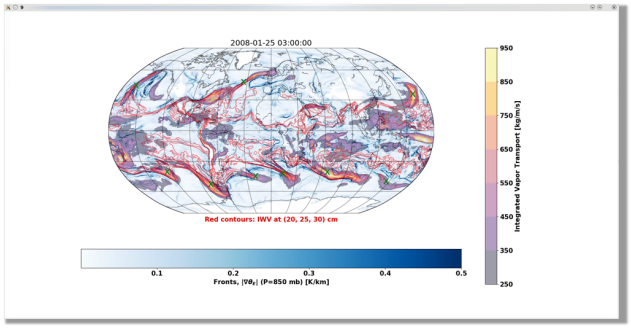

Figure 1. An example screenshot of an 3-hourly time slice of MERRA2 derived integrated vapor transport using a graphical user interface (GUI) that eight co-authors of this manuscript used to count ARs for a training dataset.

the mouse in the vicinity of an AR. A graphical indicator (a small, green 'X') is left in the location of the mouse click, which allows the user to visually assess whether they have adequately accounted for all ARs in a given field before proceeding to the next image. The GUI-relative coordinates of each click is recorded in the metadata, which allows approximate reconstruction of the geophysical location of each indicated AR. The location information is not used in constraining the Bayesian AR detection algorithm, though we do use it for understanding differences among expert contributors.

Eight of the co-authors of this manuscript (see Acknowledgements) contributed counts via this GUI, and the counts differ substantially. Each contributor counted ARs in at least 30 random time slices, with contributions ranging between 64 and 906 time slices (see Figure 2a). Figure 2a shows that the number of ARs counted varies by nearly a factor of 3 among contributors: the most 'restrictive' expert identifies a median of 4 ARs, while the most 'permissive' expert identifies a median of 11 ARs. Contributors are assigned an identification number according to the mean number of ARs counted, with the lowest 'Expert ID' (zero) having the lowest mean count and Expert ID 7 having the highest. Differences among the cumulative distributions shown in Figure 2 are mostly statistically significant, according to a suite of pair-wise Kolmogorov-Smirnov tests (Figure 2b). Counts from Expert IDs 1, 2, and 3 are mutually statistically indistinguishable at the 90% confidence level. Expert IDs 3 and 4 are likewise statistically indistinguishable, though 4 differs significantly from 1 and 2.

The differences among expert contributors leads one to wonder whether they are counting the same meteorological phenomenon, and cross-examination suggests that they are. There are a number of instances where, by chance, three experts counted ARs in the same time slice. Intercomparison of the approximate AR locations in these multiply-counted time slices (not shown) indicates that the most restrictive contributors tend to identify the same meteorological features as the most permissive contributors. The ARs identified by restrictive contributors are a subset of those identified by the permissive contributors.

These differences present two methodological challenges: (1) differences among the expert contributors will likely lead to different groups of parameter sets in a Bayesian algorithm; and (2) there is nearly an order-of-magnitude spread among the number of time slices contributed by each expert, which would lead to over-representation of the contributors with the highest number of time slices (e.g., Expert ID 5 contributed 906 counts; Figure 2a). We opt to treat all expert contributions as equally plausible, given that there is no *a priori* constraint (e.g., physical constraint or otherwise) on the number of ARs globally. Both

https://doi.org/10.5194/gmd-2020-55
Preprint. Discussion started: 17 April 2020

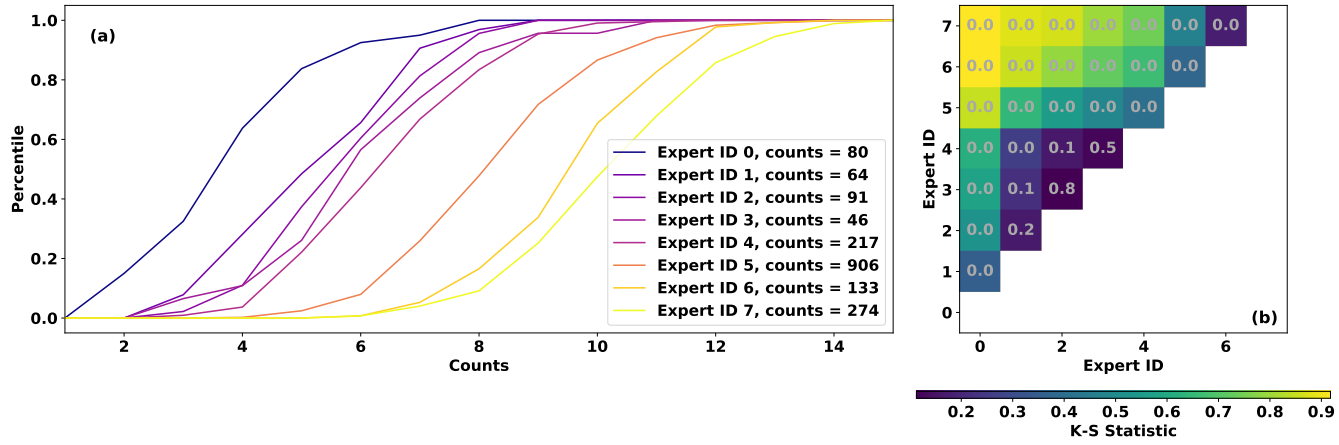

Figure 2. (a) Cumulative distributions of expert counts. (b) Two-sample Kolmogorov-Smirnov test statistics among Expert IDs. Gray text indicates the p-value; low values indicate that sets of expert counts likely have different distributions.

challenges can be addressed simply by doing the Bayesian model fitting separately for each expert and then pooling parameters in the final stage; this procedure is described in more detail in Section 2.4.1.

2.3 A Specific Implementation - TECA-BARD v1.0

We propose here a specific implementation of an AR detector on which to test the Bayesian method. For the sake of parsimony, this initial detector includes only three main criteria: contiguity above a threshold, size, and location. The detector utilizes a spatially filtered version of the IVT field, IVT' (defined toward the end of this paragraph), and in this specific implementation it seeks contiguous regions within each 2D field that are above a time-dependent threshold, where the threshold is defined as the P^{th} percentile of that specific IVT' field. This follows the motivation of Shields et al., who utilize a time-dependent threshold in order to avoid ARs becoming arbitrarily larger as water vapor mixing ratios increase in the atmosphere due to global warming. The contiguous regions must have an area that is greater than a specified threshold A_{min}. In order to avoid large contiguous regions in the tropics, associated with the intertropical convergence zone (ITCZ), the IVT field is spatially filtered as

$$\text{IVT'}(y,x) = \left(1 - e^{-2\ln 2 \cdot \frac{y^2}{\Delta y^2}}\right) \cdot \text{IVT}(y,x), \tag{3}$$

where (y,x) are spatial coordinates (latitude and longitude respectively), and Δy is the half-width at half-maximum of the filter. The filter essentially tapers the IVT field to 0 in the tropics, within a band of approximate width Δy. Table 1 summarizes the free parameters in this AR detector, and Figure 3 illustrates the stages of the detection algorithm

Table 1 also presents the prior ranges that we deem plausible for the parameter values; justification of these ranges follows. For Δy, the filter should efficiently damp the ITCZ toward 0. Though the ITCZ is relatively narrow, it migrates significantly throughout the annual cycle, so we use a minimum threshold of 5° as the lower bound. The filter should not extend so far

https://doi.org/10.5194/gmd-2020-55
Preprint. Discussion started: 17 April 2020

Geoscientific Model Development Discussions
Open Access
EGU

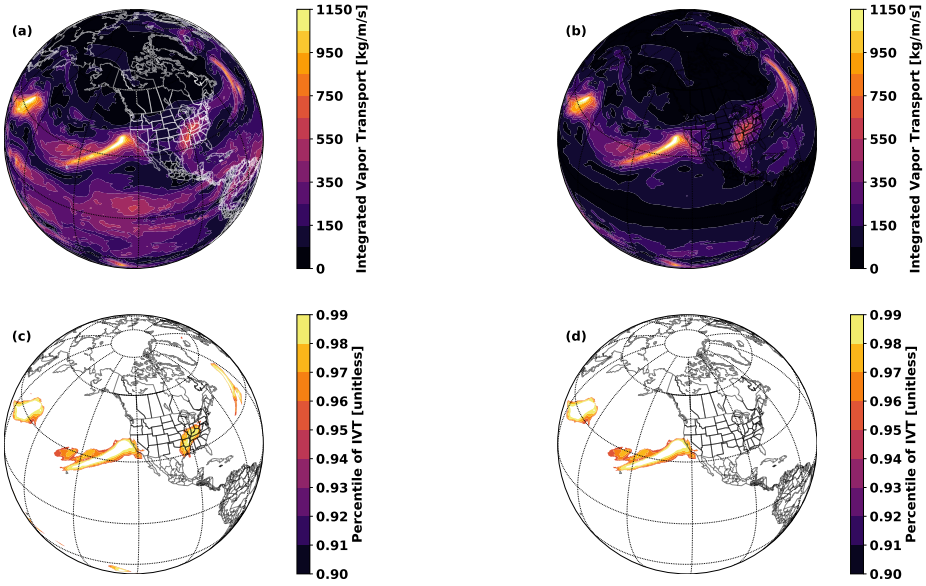

Figure 3. Illustration of the steps in TECA AR v1.0 with Δy = 15°N, P = 0.95, and A_{min} = $1 \cdot 10^{12}$ m^2: (a) the input field, integrated vapor transport (IVT); (b) IVT after application of a Δy = 15°N tropical filter (IVT'); (c) IVT' (converted to percentile) after application of the percentile filter; and (d) after application of the minimum area filter.

Table 1. Parameters, ranges, and priors in the AR detector

Parameter	Description	Range
P	Percentile threshold for IVT'	(0.8,0.99)
A_{min}	Minimum area of contiguous region	$(1 \cdot 10^{11}, 5 \cdot 10^{12})$ m^2
Δy	Zonal half-width-at-half-maximum of tropical filter	(5, 25) °N

north that it damps the midlatitudes, which is where the ARs of interest are located; hence we use an upper bound of 25°, which terminates the filter upon entering the midlatitudes. For A_{min}, we use an order-of-magnitude range based on experience in viewing ARs in meteorological data; for reference, we note that ARs are often of a size comparable to the state of California. $1{\cdot}10^{11}$ m^2 is approximately one quarter of the area of the state of California, which is likely on the too-small side, and $5{\cdot}10^{12}$ is approximately 6 times the area of California. For P, we note that the threshold is linked to the fraction of the planetary area that ARs cover in total. We use 20% of the planetary area as an upper bound ($P = 0.8$), and 0.1% as a lower bound ($P = 0.999$). The actual area covered by ARs of course depends both on the typical area of ARs and the typical number. If we assume that there are $\mathcal{O}(10)$ ARs occurring globally at any time, and they have a size $\mathcal{O}(10^{12}$ m$^2)$, then they would cover $\mathcal{O}(10\%)$ of the planetary area ($P = 0.9$) as postulated by one of the earliest AR manuscripts (Zhu and Newell, 1998).

https://doi.org/10.5194/gmd-2020-55
Preprint. Discussion started: 17 April 2020

Geoscientific Model Development Discussions

We refer to this specific implementation of AR detector, in terms of the AR counts that it yields, as F_3, such that $N_i' = F_3(P, A, \Delta y|\boldsymbol{Q_i})$. We use a half-Cauchy prior for σ, following Gelman (2006): $P_\sigma = (2/\pi s)\left(1 + (\sigma/s)^2\right)^{-1}$, and we fix the scale parameter s at a large value of 10, which permits a wide range of σ values. Given this and a choice of a uniform prior for all three parameters, and an assumption that the prior distributions are independent, this leads to a concrete Bayesian model for the posterior distribution of the parameter set $(P, A, \Delta y)$:

$$p(P, A, \Delta y \mid \boldsymbol{N}, \boldsymbol{Q}) \propto \left(\prod_{i=1}^{M} \mathcal{N}\left(N_i \mid F_3(P, A, \Delta y|\boldsymbol{Q}_i), \sigma\right) \right) \cdot \frac{2}{\pi s} \frac{1}{[1 + (\sigma/s)^2]} \tag{4}$$

2.3.1 Geometrically Constraining the Prior

The prior parameter ranges in Table 1 provide plausible prior ranges for the detector parameters, but there are some areas within this cube of parameters that we can *a priori* assert are highly improbable due to geometric considerations. This is necessary in order to avoid the Markov Chain Monte Carlo algorithm (see Section 2.4) from having points that initialize and get 'stuck' in regions of the parameter space that do not yield ARs.

By definition, the percentile threshold P will select $N_c = (1 - P) \cdot N_T$ points out of the total N_T points in the input field. If we approximate the area of all individual grid cells (ignoring for simplicity the latitudinal dependence) as $\overline{A}$, then the total area of cells above the percentile threshold will be $A_c = \overline{A} N_T (1 - P)$. By deduction, in order for any AR to be detected, the total area of grid cells above the threshold P must be as large or larger than the minimum-area threshold A for contiguous blobs above the percentile threshold: i.e., if $A_c < A$, then no AR detections are possible. We assert that parameter combinations that prohibit AR detections are implausible, and therefore the prior should be equal to 0 in such regions of parameter space. This condition effectively defines a curve in the A vs P plane, where the prior is 0 to the right of the curve:

$$A = \overline{A} N_T (1 - P)$$

This idea can be expanded further by noting that the latitude filter effectively sets values near a band $2\Delta y$ close to 0. If we assume that all points within $2\Delta y$ of the equator are effectively removed from consideration, then the total number of points under consideration N_T should be reduced by the fraction f of points that are taken out by the filter. In the latitudinal direction, cell areas are only a function of latitude y ($\cos(y)$ specifically), so with the above assumption, f can be approximated simply as:

$$f = 1 - \frac{\int_{-\Delta y}^{\Delta y} \cos(y)\, dy}{\int_{-\pi/2}^{\pi/2} \cos(y)\, dy} = 1 - \sin(\Delta y).$$

With this, the number of cells picked out by the filter will be approximately $N_c' = f \cdot N_t \cdot (1 - P)$. If we assume that there are $\mathcal{O}(10)$ ARs at any given time, then there are typically at most $N_c'/10$ grid cells per AR. If we tighten the constraint to assert

https://doi.org/10.5194/gmd-2020-55
Preprint. Discussion started: 17 April 2020


Geoscientific Model Development Discussions

that these conditions should lead to ARs that typically have more than 1 grid cell per AR, then this leads to a formulation of the prior constraint that depends on the value of the latitude filter:

$$A = \overline{A} \cdot (1 - \sin(2\Delta y)) \cdot \frac{N_T \cdot (1 - P)}{10} \tag{5}$$

We modify the uniform prior to be equal to 0 outside the surface defined in Equation 5.

2.4 Markov Chain Monte Carlo Sampling

We use an affine-transformation-invariant Markov Chain Monte Carlo (MCMC) sampling method (Goodman and Weare, 2010), implemented in Python by Foreman-Mackey et al. (2013) (`emcee v2.1.1`[1]), to approximately sample from the posterior distribution described in Equation 4. We utilize 1,024 MCMC 'walkers' (semi-independent MCMC chains) with starting positions sampled uniformly from the parameter ranges shown in Table 1. Parameter values outside the parameter-surface described by Equation 5 are rejected and randomly sampled until all initial parameter sets satisfy Equation 5.

The MCMC algorithm essentially finds sets of parameters for which TECA-BARD yields sets of AR counts that are close (in a least-squares sense) to the input set of expert counts described in Section 2.2. Within an MCMC step, each walker proposes a new set of parameters. Each MCMC walker runs the TECA-BARD algorithm described in Section 2.3, for its set of proposed parameter values, on the IVT field (Q_i) from all time slices in MERRA2 for which there are expert counts N_i; TECA-BARD (F_3 in Equation 4) returns the global number of ARs N_i' for each time slice. The sets of expert counts and TECA-BARD counts are provided as input to Equation 2, which is then used in Equation 1 to evaluate the posterior probability of the proposed parameters. The proposed parameter is then either accepted or rejected following the algorithm outlined by Foreman-Mackey et al. (2013). Parameters with higher posterior probabilities generally have a higher chance of being accepted. The accept/reject step has an adjustable parameter (a in Equation 10 of Foreman-Mackey et al., 2013), which we set to a value of 2, following Goodman and Weare (2010). Sensitivity tests with this value showed little qualitative change in the output of the MCMC samples.

We run all 1,024 MCMC walkers for 1,000 steps and extract MCMC samples from the last step. Manual examination of the evolution of parameters within individual walker chains shows that the evolution of parameters reaches a statistical equilibrium after $\mathcal{O}(100)$ steps, so the chains should all be well-equilibrated by 1,000 steps. We ran a brute-force calculation of the posterior distribution on a regularly-spaced grid of parameter values (not shown) to verify that the MCMC algorithm is indeed sampling correctly from the posterior distribution.

https://doi.org/10.5194/gmd-2020-55
Preprint. Discussion started: 17 April 2020

Geoscientific Model Development Discussions

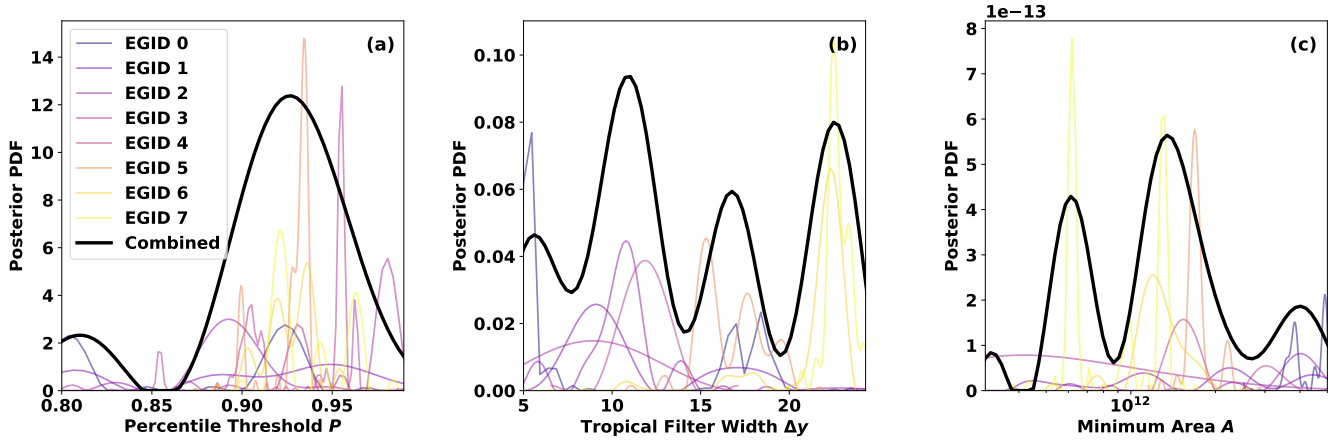

Figure 4. Posterior marginal distributions of parameters for each Expert Group ID (EGID) $p_i(\boldsymbol{\theta})$ (colored curves) and for the full, combined posterior $p(\boldsymbol{\theta})$ (black curve): (a) percentile threshold for IVT' P, (b) zonal half-width-at-half-maximum of the tropical filter Δy, and (c) minimum area of contiguous regions A_{min}. Posterior distributions for EGIDs are scaled by $\frac{1}{8}$, consistent with Equation 6

2.4.1 Expert Groups

In order to address the challenges posed by having AR count datasets that differ significantly among expert contributors (described in Section 2.2), we develop a separate posterior model for each Expert ID j: $p_j(\boldsymbol{\theta} \mid \boldsymbol{N}_j, \boldsymbol{Q})$. The final model is a normalized, unweighted sum of posterior distributions from each Expert ID:

$$p(\boldsymbol{\theta} \mid \boldsymbol{N}, \boldsymbol{Q}) = \frac{1}{8} \cdot [p_0(\boldsymbol{\theta} \mid \boldsymbol{N}_0, \boldsymbol{Q}) + p_1(\boldsymbol{\theta} \mid \boldsymbol{N}_1, \boldsymbol{Q}) + \ldots + p_7(\boldsymbol{\theta} \mid \boldsymbol{N}_7, \boldsymbol{Q})] \tag{6}$$

Practically speaking, we achieve this by running the MCMC integration separately for each Expert ID and then combining the MCMC samples together. TECA BARD v1.0 uses 128 samples from each Expert ID, for a total of 1,024 samples. The samples are stored in an input parameter table such that parameters from the same Expert ID are contiguous, which allows post hoc grouping of results by Expert ID. We refer to these groups by their *Expert Group IDs*, which correspond to data from each Expert ID used in the MCMC integration. Figure 4 shows marginal distributions of the TECA-BARD v1.0 parameters.

2.5 Implementation in the Toolkit for Extreme Climate Analysis

We implement the detector as an application in the Toolkit for Extreme Climate Analysis (TECA[2]). TECA is a framework for facilitating parallel analysis of petabyte-scale datasets. TECA provides generic modular components that implement parallel execution patterns and scalable I/O. These components can easily be composed into analysis pipelines that run efficiently at

[1]https://github.com/dfm/emcee/releases/tag/v2.2.1
[2]https://github.com/lbl-eesa/teca doi:10.20358/C8C651

https://doi.org/10.5194/gmd-2020-55
Preprint. Discussion started: 17 April 2020

Geoscientific Model Development Discussions
Open Access
EGU

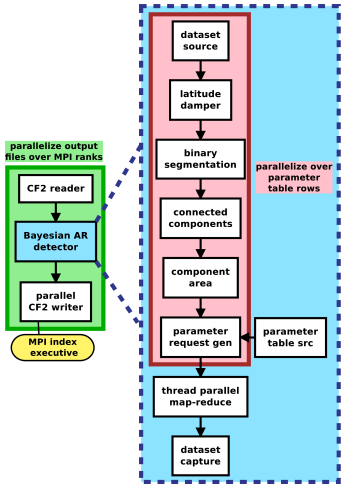

Figure 5. A diagram of the TECA pipeline that makes up the TECA Bayesian AR Detector application.

scale at high performance computing (HPC) centers. Figure 5 shows the modular components used to compose the TECA-BARD v1.0 application. TECA is primarily written in C++, and it offers Python bindings to facilitate prototyping of pipelines in a commonly-used scientific language. Early versions of TECA-BARD v1.0 were developed using these bindings, and the MCMC code invokes TECA-BARD via these Python bindings.

To improve performance on large calculations, TECA uses a map-reduce framework that takes advantage of both thread-level parallelism (using C++ threads) and multi-core parallelism (with the message passing interface MPI). TECA-BARD v1.0 distributes a range of MCMC parameters over different threads, and it distributes timesteps over different processes using MPI. This strategy allows TECA-BARD v1.0 to scale efficiently on HPC systems. We ran TECA-BARD v1.0, which effectively consists of 1,024 separate AR detectors, on 36.5 years of 3-hourly MERRA2 output (see Section 3) at the National Energy Research Scientific Computing center (NERSC) on the Cori system using 396 68-core Intel Xeon-Phi (Knights Landing) nodes in approximately 4 minutes (wallclock time).

3 Evaluation of TECA-BARD v1.0

We run TECA-BARD v1.0 on 3-hourly IVT output from the MERRA2 reanalysis (Gelaro et al., 2017) from 1 January, 1980 to 30 June 2017, which involves running the detector described in Section 2 for each of the 1,024 samples from the posterior distribution. Output from TECA-BARD v1.0 differs in character from other algorithm output in ARTMIP in that it provides a posterior probability of AR detection p_{AR}, rather than a binary indicator of AR presence (Shields et al., 2018). We derive a comparable measure of AR presence by averaging binary AR identifications across available ARTMIP algorithms, on a location-by-location basis. This yields a probability-like quantity, which we refer to as the 'ARTMIP Confidence Index',

https://doi.org/10.5194/gmd-2020-55
Preprint. Discussion started: 17 April 2020

Geoscientific Model Development Discussions
Open Access
EGU

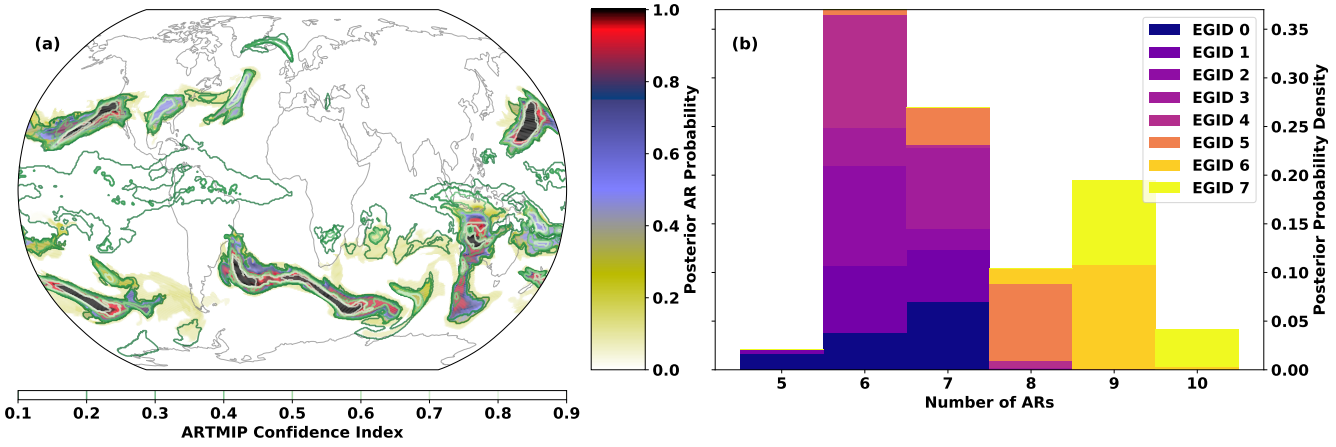

Figure 6. (a) AR detections on 7 February, 2019 at 0600Z. Filled contours show the posterior probability of AR detection p_{AR} from TECA-BARD v1.0. Contour lines show the ARTMIP Confidence Index (the proportion of ARTMIP algorithms detecting an AR at a given location), P_{ARTMIP}. (b) Posterior distributions of counts for the same time slice, grouped by Expert Group ID (EGID).

P_{ARTMIP}: the proportion of ARTMIP algorithms reporting AR presence at each time slice. Output from TECA-BARD v1.0 is shown in Figure 6a, which also shows the corresponding ARTMIP Confidence Index for comparison.

TECA-BARD v1.0 and ARTMIP generally agree on the presence of 'high confidence' ARs: regions in which p_{AR} and P_{ARTMIP} are high. There are five regions of extremely high posterior AR probability in TECA-BARD-v1.0: areas where $p_{\text{AR}} \approx 1.0$ (regions with black coloring) in Figure 6a. All five of these regions are enclosed by white contours, indicating that at least 90% of ARTMIP algorithms also indicate AR presence. There are two additional distinct regions (in the the eastern U.S. and the central, north Atlantic) where $P_{\text{ARTMIP}} > 0.9$, whereas p_{AR} only reaches approximately 0.6; these regions have relatively small areas. Such behavior arises because of multimodality in the posterior distribution of parameters; e.g., Figure 4c shows that there are several distinct modes in the minimum area parameter. It is likely that these two regions of high IVT have areas that fall between two of these modes.

Most of the disagreement between ARTMIP and TECA-BARD v1.0 is associated with 'low confidence' AR regions: particularly regions in which the ARTMIP Confidence Index is in the range of 20%. The most prominent of these is a large region of $P_{\text{ARTMIP}} \approx 0.2$ in the tropics; whereas $p_{\text{AR}} \approx 0$ throughout the tropics. We argue that this represents erroneous detection of the ITCZ by a small subset of ARTMIP algorithms. The tropical filter (corresponding to parameter Δy) in TECA-BARD v1.0 explicitly filters out the tropics to avoid such erroneous detection of the ITCZ.

Figure 6b shows that TECA-BARD v1.0 detects 5–10 ARs in the dataset shown in Figure 6a. The range of uncertainty is much smaller within individual Expert Group IDs (EGIDs); the most restrictive Expert Group ID (EGID 0) detects 5–7 ARs, whereas the most permissive parameter group (EGID 7) detects 9–10 ARs. This is consistent with the behavior shown in Figure2a; lower Expert IDs have lower counts and vice-versa.

https://doi.org/10.5194/gmd-2020-55
Preprint. Discussion started: 17 April 2020

Geoscientific Model Development Discussions
Open Access
EGU

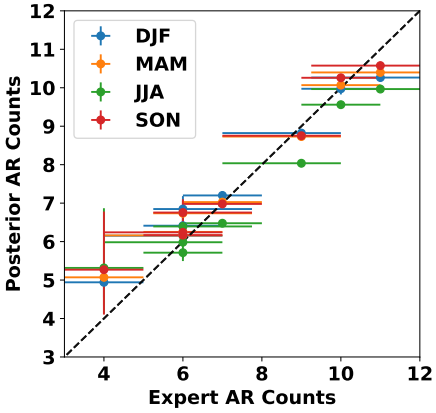

Figure 7. Posterior mean AR counts for each season, grouped by expert group ID vs median number of ARs counted by expert contributors. Whiskers indicate the 5-95 percentile range. The dashed line shows the 1:1 line.

More broadly, the number of ARs counted within each EGID is consistent with the number of ARs counted by the corresponding expert contributors. Figure 7 shows that the AR counts from the various EGIDs are consistent with AR count statistics from the corresponding expert contributors. For all seasons, the points in Figure 7 are close–within error bars–to the one-to-one line. Note that we do not disaggregate expert counts by season, since doing so would lead to small sample sizes for some expert IDs. The seasonal range in posterior counts across EGIDs suggests that this should not affect our conclusion that EGIDs within TECA-BARD v1.0 emulate the counting statistics of corresponding experts, since the seasonal range is only approximately $\pm$ 1.

Figure 7 also shows that the uncertainty in the number of detected ARs in TECA-BARD v1.0 is a direct consequence of uncertainty in the input dataset. Further, the spread in expert counts results in EGIDs having distinct groups of parameters. Figure 4 shows that the EGIDs associated with the most restrictive experts tend to have large minimum area parameters and narrower tropical filters, whereas the opposite is true for the most permissive EGIDs. This shows that the MCMC method yields a set of parameters that yield AR detectors that emulate the bulk counting statistics of the input data.

4 Uncertainty in the Relationship between ENSO and AR Count

We assess the impact of parametric uncertainty in TECA-BARD v1.0 by asking a relatively simple question: *Are there more ARs during El Niño events?*. We examine this question from a global perspective, which is partly motivated by Guan and Waliser (2015) (their Figures 10a,b), who show coherent changes in AR probability associated with the El Niño Southern Oscillation (ENSO). The predominant effect is an equatorward shift of ARs during the positive phase of ENSO, and their figure seems to show more areas of increased AR occurrence than areas of decrease; this might suggest that positive phases of ENSO are associated with ARs globally. Goldenson et al. (2018) indicate that, at least regionally, their analysis of the impact of ENSO on AR predictability leads to a different conclusion than that of Guan and Waliser (2015). Goldenson et al. (2018) and

https://doi.org/10.5194/gmd-2020-55
Preprint. Discussion started: 17 April 2020

Geoscientific Model Development Discussions
Open Access
EGU

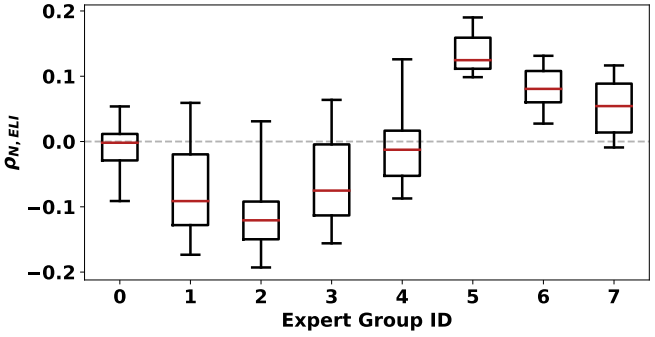

Figure 8. Box-and-whisker plots of the correlation between global AR count and ELI $\rho_{N,\mathrm{ELI}}$, grouped by EGID. Red lines show the median, box limits show the interquartile range, and whiskers indicate the 5-95% range.

Guan and Waliser (2015) utilize different AR detection algorithms, which suggests that inferred relationships between ENSO and ARs may depend on the detection algorithm used.

TECA-BARD v1.0 consists of 1,024 plausible AR detectors, which allows us to analyze whether there are significant differences in the answer to this question across the sets of AR detector parameters. We compare the TECA-BARD v1.0 output from MERRA2, described in Section 3, against the ENSO Longitude Index (ELI) of Williams and Patricola (2018). We calculate the average ELI for each boreal winter (December, January, and February) between 1981 and 2017. Similarly, we calculate the DJF-average number of detected ARs, over the same time period, for each of the 1,024 sets of parameters in TECA-BARD v1.0; we then calculate the Spearman rank correlation coefficient $\rho_{N,\mathrm{ELI}}$ between each set of DJF AR counts and ELI. This yields 1,024 values of $\rho_{N,\mathrm{ELI}}$, which expresses the interannual correlation between DJF AR count and ELI for each set of AR detectors in TECA-BARD v1.0. Figure 8 shows the results of this calculation.

Across all EGIDs, the correlation coefficients range from approximately -0.2 to +0.2; they span zero. However, grouping results by EGID shows that different groups of detector parameters yield qualitatively different results. Figure 8 shows the posterior statistics of $\rho_{N,\mathrm{ELI}}$, grouped by EGID. EGIDs 0-4 have predominantly negative correlation coefficients (the medians and 5$^{\mathrm{th}}$ percentile values are all negative), though the 95$^{\mathrm{th}}$ percentile values are all positive. On the other hand, correlation coefficients from EGIDs 5 and 6 are all entirely positive, and most values from EGID 7 are positive. Even within the uncertainty quantification framework of TECA-BARD v1.0, if we had utilized a single expert contributor–e.g., EGID 6 or 7–we might have over-confidently concluded that there are more ARs globally during El Niño events.

It is intriguing that the most restrictive EGIDs tend to yield negative correlation coefficients, while the most permissive EGIDs tend to yield positive correlation coefficients. This variation appears to be predominantly controlled by variations in the tropical filter Δy. Figures 9 show samples of the posterior distribution of $\rho_{N,\mathrm{ELI}}$ as a function of detector parameters. In Figures 9a,c, $\rho_{N,\mathrm{ELI}}$ is evenly distributed across zero for the entire parameter space; only in Figure 9b does the correlation coefficient show systematic variation with the input parameter Δy.

https://doi.org/10.5194/gmd-2020-55
Preprint. Discussion started: 17 April 2020

Geoscientific Model Development Discussions

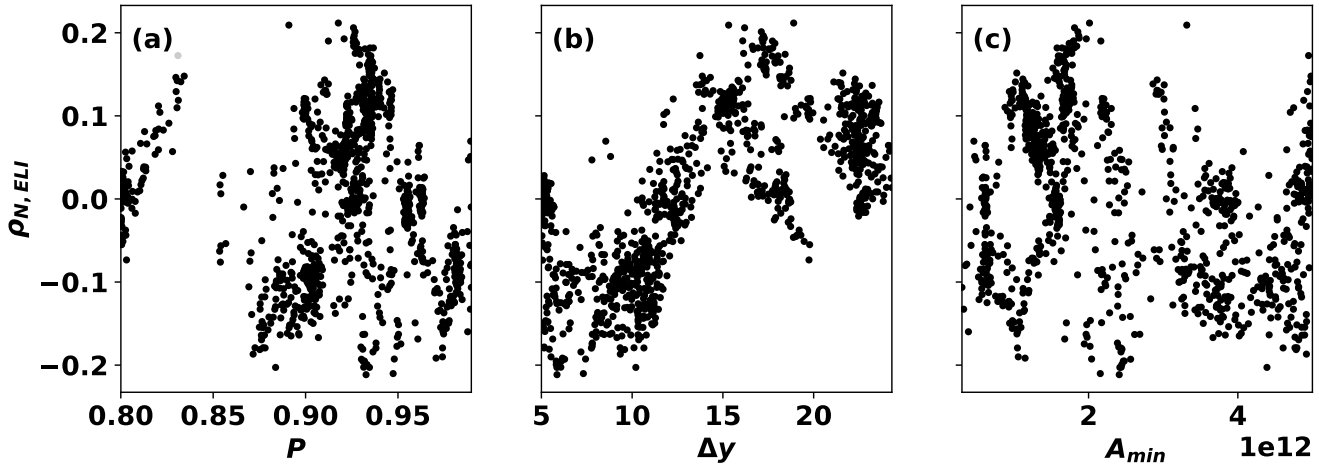

Figure 9. Correlation between global AR count and ELI $\rho_{N,\mathrm{ELI}}$ as a function of AR detector parameter.

We further disaggregate results in Figure 10 by showing how $\rho_{N,\mathrm{ELI}}$ clusters by EGID in two-dimensional projections of the parameter space. We utilize `fastKDE`[3] (O'Brien et al., 2016) to calculate two-dimensional marginal posterior distributions for each EGID: e.g., $p_j(P, \Delta y \mid \boldsymbol{N}, \boldsymbol{Q})$ in Figure 10a (where j corresponds to the EGID). We show contours of constant p_j, colored by EGID, such that 95% of the posterior distribution for each EGID falls within the given contour; the colored contours in Figure 10 effectively outline the parameter samples for each EGID.

Parameter clusters with both positive $\rho_{N,\mathrm{ELI}}$ and high Δy tend to form distinct zones of points in Figure 10: clusters with relatively low A_{min} and relatively high Δy. Parameters with negative $\rho_{N,\mathrm{ELI}}$ predominantly fall along a line in the P-A_{min} plane in Figure 10b, with the positive $\rho_{N,\mathrm{ELI}}$ values forming separate clusters of points off the line. These separate clusters are associated with the more permissive EGIDs.

We argue that the differences in correlation coefficient between the restrictive and permissive EGIDs likely results from differences in the degree to which tropical moisture anomalies are filtered among the EGIDs. Patricola et al. (2019) show that strong El Niño events are associated with positive IVT anomalies in much of the tropics and a separate band of positive anomalies in the midlatitudes (around 30° latitude; their Figure 11). The positive IVT anomalies in the tropics would have no effect on the subset of AR detector parameters with high values of Δy, since these values would be aggressively filtered. This subset of parameters with high Δy–which is associated with the permissive EGIDs and positive values of $\rho_{N,\mathrm{ELI}}$ (Figures 4b and 9b)–would then only be affected by the higher-than-average IVT in the midlatitudes. This would result in larger numbers of ARs during El Niño events. For AR detectors parameters with low values of Δy, the zone of positive anomaly in the tropics would not be totally filtered out, which increases the chances for zones of high IVT in the midlatitudes to be connected to zones of high IVT in the tropics. This could potentially result in larger-than-average, and fewer, ARs during El Niño.

[3]https://bitbucket.org/lbl-cascade/fastkde at commit f2564d6

https://doi.org/10.5194/gmd-2020-55
Preprint. Discussion started: 17 April 2020

Geoscientific Model Development Discussions

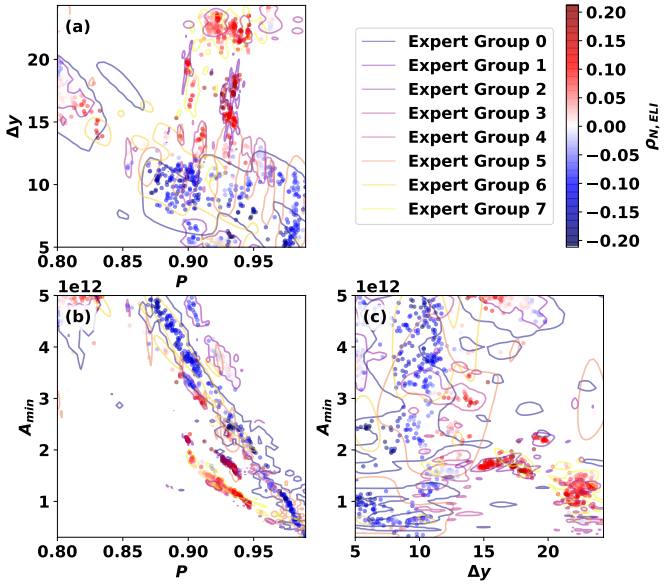

Figure 10. Pairplots of AR detector samples. Marker colors indicate the correlation between global AR count and ELI $\rho_{N,\mathrm{ELI}}$, and colored lines show contours of constant p_j such that 95% of the posterior distribution for each EGID falls within the given contour.

5 The Importance of Uncertainty in Feature Detection

The results in Section 4 show that equally plausible sets of AR detector parameters can yield qualitatively different conclusions about the connection between ENSO and AR count. These results also show that the data used to constrain the AR detector parameters in TECA-BARD v1.0 has a huge influence on the choice of parameters and ultimately the conclusions that one might draw. Figure 8 shows that almost half of the spread in $\rho_{N,\mathrm{ELI}}$ can be explained by the spread in expert counts used to constrain the Bayesian model. This spread results from differences in subjective opinion about what does or does not constitute an atmospheric river.

In the current literature, AR detectors have two main developmental stages: (1) decide on the steps used in the AR detection algorithm, and (2) determine values used for unconstrained parameters (e.g., thresholds like P, Δy, and A_{min}). In all examples of AR literature known to these authors, both steps rely on expert judgement. If we frame this in terms of the AR detector described in Section 2.3, step (2) would involve an expert varying the detector parameters P, Δy, and A_{min} until the resulting AR detections are acceptable. It seems reasonable to assume that if Expert ID 3 were to manually choose parameters in such a way, they would likely choose parameters that would yield a negative correlation between ENSO and global AR count; conversely, Expert ID 6 would almost certainly choose parameters that would yield a positive correlation coefficient. Setting aside uncertainty in the detector design (stage 1), two different experts could potentially develop AR detectors that would come to opposite conclusions about the impact of ENSO on AR count.

https://doi.org/10.5194/gmd-2020-55
Preprint. Discussion started: 17 April 2020


Geoscientific Model Development Discussions

It is crucial to recognize the importance and impact of this spread in subjective opinion. Subjective opinion is currently used in the literature to define quantitative methods for detecting ARs. Since we currently lack physical theories to constrain AR detection schemes like this, such as theories about what the number of ARs should be, subjective opinion is the only option. These results show that subjective opinion can qualitatively impact the conclusions that one might draw.

This study considers the parametric uncertainty in a single detector framework, and it does not consider the structural uncertainty in the detector framework itself. This is a key limitation of this study, and it is an opportunity for expanding this work in future studies. For example, we could have utilized an absolute threshold in IVT (e.g., 250 kg m^{-1} s^{-1}) rather than a relative, percentile-based threshold. One might imagine applying the general Bayesian framework described in Section 2.1 to other existing AR detectors in the literature as a way to explore both structural and parametric uncertainty. The expert count data produced as part of this study, which are publicly available following information in the *Code and data availability* statement at the end of this manuscript, could readily be used for such an exercise.

The approach described here could also be applied to detectors of other types of weather phenomena. For example, the U.S. Clivar Hurricane Working Group determined that some tropical cyclone research results depend on how tropical cyclones are detected: particularly results concerning weaker cyclones (Walsh et al., 2015). Similarly, the Intercomparison of Mid Latitude Storm Diagnostics (IMILAST) project determined that scientific results regarding extratropical cyclones can depend on how they are detected (Neu et al., 2013). There is also emerging research on frontal systems that could be interpreted to suggest a similar uncertainty with respect to tracking method (Schemm et al., 2018). We argue that such uncertainty is inherent to heuristic phenomena detectors, and Bayesian approaches like the one described in Section 2.1 could be used to quantify this uncertainty.

Code and data availability.

The code for TECA-BARD v1.0 is available at https://github.com/LBL-EESA/TECA under TECA release 'TECA-BARD-V1.0'. TECA-BARD v1.0 is available as a TECA application `teca_ar_detect` under source file `apps/teca_ar_detect.cxx`. The code for sampling the posterior distribution of the TECA-BARD parameters is available at https://bitbucket.org/lbl-cascade/bayesian_ar_detector. Data containing the AR counts used for constraining the TECA-BARD parameters are available at https://portal.nersc.gov/archive/home/projects/cascade/www/teca_bard/ar_count_files.tar and data containing posterior distribution samples are available at https://doi.org/10.5281/zenodo.3677543.

Author contributions. TAO was responsible for development of the concept, development of the statistical method, implementation of method, generation of figures, and initial drafting of the manuscript. MDR contributed to the development of the statistical method. TAO, BL, AAE, HK, JNJ, and Prabhat all contributed to the implementation of the method in TECA. TAO, CMP, JPO, AM, SAR, AMR, AC, and HAI contributed to the database of AR counts. All authors contributed to the editing of the manuscript.

https://doi.org/10.5194/gmd-2020-55
Preprint. Discussion started: 17 April 2020

Geoscientific Model Development Discussions

Competing interests. The authors declare they have no conflict of interest.

Acknowledgements. This research was supported by the Director, Office of Science, Office of Biological and Environmental Research of the U.S. Department of Energy Regional and Global Climate Modeling Program (RGCM) and used resources of the National Energy Research Scientific Computing Center (NERSC), also supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The authors thank Dr. Christopher J. Paciorek for providing useful input on the manuscript.

https://doi.org/10.5194/gmd-2020-55
Preprint. Discussion started: 17 April 2020

Geoscientific Model Development Discussions

References

Dong, L., Leung, L. R., Song, F., and Lu, J.: Roles of SST versus internal atmospheric variability in winter extreme precipitation variability along the U.S. West Coast, Journal of Climate, pp. JCLI–D–18–0062.1, https://doi.org/10.1175/JCLI-D-18-0062.1, http://journals.ametsoc.org/doi/10.1175/JCLI-D-18-0062.1, 2018.

Foreman-Mackey, D., Hogg, D. W., Lang, D., and Goodman, J.: emcee : The MCMC Hammer, Publications of the Astronomical Society of the Pacific, 125, 306–312, https://doi.org/10.1086/670067, http://arxiv.org/abs/1202.3665{%}0Ahttp://dx.doi.org/10.1086/670067http://iopscience.iop.org/article/10.1086/670067, 2013.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G. K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), Journal of Climate, 30, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1, 2017.

Gelman, A.: Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper), Bayesian Analysis, 1, 515–534, https://doi.org/10.1214/06-BA117A, http://projecteuclid.org/euclid.ba/1340371048, 2006.

Gershunov, A., Shulgina, T., Clemesha, R. E. S., Guirguis, K., Pierce, D. W., Dettinger, M. D., Lavers, D. A., Cayan, D. R., Polade, S. D., Kalansky, J., and Ralph, F. M.: Precipitation regime change in Western North America: The role of Atmospheric Rivers, Scientific Reports, 9, 9944, https://doi.org/10.1038/s41598-019-46169-w, https://doi.org/10.1038/s41598-019-46169-w.http://www.nature.com/articles/s41598-019-46169-w, 2019.

Goldenson, N., Leung, L. R., Bitz, C. M., and Blanchard-Wrigglesworth, E.: Influence of Atmospheric Rivers on Mountain Snowpack in the Western United States, Journal of Climate, 31, 9921–9940, https://doi.org/10.1175/JCLI-D-18-0268.1, http://journals.ametsoc.org/doi/10.1175/JCLI-D-18-0268.1, 2018.

Goodman, J. and Weare, J.: Ensemble samplers with affine invariance, Communications in Applied Mathematics and Computational Science, 5, 65–80, https://doi.org/10.2140/camcos.2010.5.65, http://msp.org/camcos/2010/5-1/p04.xhtml, 2010.

Gorodetskaya, I. V., Tsukernik, M., Claes, K., Ralph, M. F., Neff, W. D., and Van Lipzig, N. P. M.: The role of atmospheric rivers in anomalous snow accumulation in East Antarctica, Geophysical Research Letters, 41, 6199–6206, https://doi.org/10.1002/2014GL060881, http://doi.wiley.com/10.1002/2014GL060881, 2014.

Guan, B. and Waliser, D. E.: Detection of atmospheric rivers: Evaluation and application of an algorithm for global studies, Journal of Geophysical Research: Atmospheres, 120, 12 514–12 535, https://doi.org/10.1002/2015JD024257, http://doi.wiley.com/10.1002/2015JD024257, 2015.

Hu, H., Dominguez, F., Wang, Z., Lavers, D. A., Zhang, G., and Ralph, F. M.: Linking Atmospheric River Hydrological Impacts on the U.S. West Coast to Rossby Wave Breaking, Journal of Climate, 30, 3381–3399, https://doi.org/10.1175/JCLI-D-16-0386.1, https://doi.org/10.1175/JCLI-D-16-0386.1http://journals.ametsoc.org/doi/10.1175/JCLI-D-16-0386.1, 2017.

Huning, L. S., Margulis, S. A., Guan, B., Waliser, D. E., and Neiman, P. J.: Implications of Detection Methods on Characterizing Atmospheric River Contribution to Seasonal Snowfall Across Sierra Nevada, USA, Geophysical Research Letters, 44, 10,445–10,453, https://doi.org/10.1002/2017GL075201, http://doi.wiley.com/10.1002/2017GL075201, 2017.

https://doi.org/10.5194/gmd-2020-55
Preprint. Discussion started: 17 April 2020

Huning, L. S., Guan, B., Waliser, D. E., and Lettenmaier, D. P.: Sensitivity of Seasonal Snowfall Attribution to Atmospheric Rivers and Their Reanalysis-Based Detection, Geophysical Research Letters, 46, 794–803, https://doi.org/10.1029/2018GL080783, http://doi.wiley.com/10.1029/2018GL080783, 2019.

Kunkel, K. E., Easterling, D. R., Kristovich, D. A. R., Gleason, B., Stoecker, L., and Smith, R.: Meteorological Causes of the Secular Variations in Observed Extreme Precipitation Events for the Conterminous United States, Journal of Hydrometeorology, 13, 1131–1141, https://doi.org/10.1175/JHM-D-11-0108.1, http://journals.ametsoc.org/doi/abs/10.1175/JHM-D-11-0108.1, 2012.

Neu, U., Akperov, M., Bellenbaum, N., Benestad, R., Blender, R., Caballero, R., Cocozza, A., Dacre, H., Feng, Y., Fraedrich, K., Grieger, J., Gulev, S., Hanley, J., Hewson, T., Inatsu, M., Keay, K., Kew, S., Kindem, I., Leckebusch, G. C., Liberato, M., Lionello, P., Mokhov, I., Pinto, J., Raible, C., Reale, M., Rudeva, I., Schuster, M., Simmonds, I., Sinclair, M., Sprenger, M., Tilinina, N., Trigo, I., Ulbrich, S., Ulbrich, U., Wang, X., and Wernli, H.: IMILAST: A Community Effort to Intercompare Extratropical Cyclone Detection and Tracking Algorithms, Bulletin of the American Meteorological Society, 94, 529–547, https://doi.org/10.1175/BAMS-D-11-00154.1, 2013.

O'Brien, T. A., Kashinath, K., Cavanaugh, N. R., Collins, W. D., and O'Brien, J. P.: A fast and objective multidimensional kernel density estimation method: FastKDE, Computational Statistics and Data Analysis, 101, 148–160, https://doi.org/10.1016/j.csda.2016.02.014, http://linkinghub.elsevier.com/retrieve/pii/S0167947316300408, 2016.

Patricola, C. M., O'Brien, J. P., Risser, M. D., Rhoades, A. M., O'Brien, T. A., Ullrich, P. A., Stone, D. A., and Collins, W. D.: Maximizing ENSO as a Source of Western US Hydroclimate Predictability, Climate Dynamics, In Review, 2019.

Ralph, F. M., Dettinger, M. D., Cairns, M. M., Galarneau, T. J., and Eylander, J.: Defining "Atmospheric River": How the Glossary of Meteorology Helped Resolve a Debate, Bulletin of the American Meteorological Society, 99, 837–839, https://doi.org/10.1175/BAMS-D-17-0157.1, https://doi.org/10.1175/BAMS-D-17-0157.1http://journals.ametsoc.org/doi/10.1175/BAMS-D-17-0157.1, 2018.

Ralph, F. M., Rutz, J. J., Cordeira, J. M., Dettinger, M., Anderson, M., Reynolds, D., Schick, L. J., and Smallcomb, C.: A Scale to Characterize the Strength and Impacts of Atmospheric Rivers, Bulletin of the American Meteorological Society, 100, 269–289, https://doi.org/10.1175/BAMS-D-18-0023.1, https://doi.org/10.1175/BAMS-D-18-0023.1http://journals.ametsoc.org/doi/10.1175/BAMS-D-18-0023.1, 2019.

Rutz, J. J., Shields, C. A., Lora, J. M., Payne, A. E., Guan, B., Ullrich, P., O'Brien, T., Leung, L. R., Ralph, F. M., Wehner, M., Brands, S., Collow, A., Goldenson, N., Gorodetskaya, I., Griffith, H., Kashinath, K., Kawzenuk, B., Krishnan, H., Kurlin, V., Lavers, D., Magnusdottir, G., Mahoney, K., McClenny, E., Muszynski, G., Nguyen, P. D., Prabhat, Qian, Y., Ramos, A. M., Sarangi, C., Sellars, S., Shulgina, T., Tome, R., Waliser, D., Walton, D., Wick, G., Wilson, A. M., and Viale, M.: The Atmospheric River Tracking Method Intercomparison Project (ARTMIP): Quantifying Uncertainties in Atmospheric River Climatology, Journal of Geophysical Research: Atmospheres, 53, 2019JD030 936, https://doi.org/10.1029/2019JD030936, https://onlinelibrary.wiley.com/doi/abs/10.1029/2019JD030936, 2019.

Schemm, S., Sprenger, M., and Wernli, H.: When during Their Life Cycle Are Extratropical Cyclones Attended by Fronts?, Bulletin of the American Meteorological Society, 99, 149–165, https://doi.org/10.1175/BAMS-D-16-0261.1, http://journals.ametsoc.org/doi/10.1175/BAMS-D-16-0261.1, 2018.

Sellars, S. L., Kawzenuk, B., Nguyen, P., Ralph, F. M., and Sorooshian, S.: Genesis, Pathways, and Terminations of Intense Global Water Vapor Transport in Association with Large-Scale Climate Patterns, Geophysical Research Letters, 44, 12,465–12,475, https://doi.org/10.1002/2017GL075495, 2017.

Shields, C. A., Rutz, J. J., Leung, L.-Y., Ralph, F. M., Wehner, M., Kawzenuk, B., Lora, J. M., McClenny, E., Osborne, T., Payne, A. E., Ullrich, P., Gershunov, A., Goldenson, N., Guan, B., Qian, Y., Ramos, A. M., Sarangi, C., Sellars, S., Gorodetskaya, I., Kashinath, K., Kurlin, V., Mahoney, K., Muszynski, G., Pierce, R., Subramanian, A. C., Tome, R., Waliser, D., Walton, D., Wick, G., Wilson, A., Lavers,

https://doi.org/10.5194/gmd-2020-55
Preprint. Discussion started: 17 April 2020

D., Collow, A., Krishnan, H., Magnusdottir, G., and Nguyen, P.: Atmospheric River Tracking Method Intercomparison Project (ARTMIP): project goals and experimental design, Geoscientific Model Development, 11, 2455–2474, https://doi.org/10.5194/gmd-11-2455-2018, https://www.geosci-model-dev.net/11/2455/2018/, 2018.

Walsh, K., Camargo, S., Vecchi, G., Daloz, A., Elsner, J., Emanuel, K., Horn, M., Lim, Y.-K., Roberts, M., Patricola, C., Scoccimarro, E., Sobel, A., Strazzo, S., Villarini, G., Wehner, M., Zhao, M., Kossin, J., LaRow, T., Oouchi, K., Schubert, S., Wang, H., Bacmeister, J., Chang, P., Chauvin, F., Jablonowski, C., Kumar, A., Murakami, H., Ose, T., Reed, K., Saravanan, R., Yamada, Y., Zarzycki, C., Vidale, P., Jonas, J., and Henderson, N.: Hurricanes and Climate: The U.S. CLIVAR Working Group on Hurricanes, Bulletin of the American Meteorological Society, 96, 997–1017, https://doi.org/10.1175/BAMS-D-13-00242.1, 2015.

Wehner, M. F., Reed, K. A., Loring, B., Stone, D., and Krishnan, H.: Changes in tropical cyclones under stabilized 1.5 and 2.0 °C global warming scenarios as simulated by the Community Atmospheric Model under the HAPPI protocols, Earth System Dynamics, 9, 187–195, https://doi.org/10.5194/esd-9-187-2018, https://www.earth-syst-dynam.net/9/187/2018/, 2018.

Williams, I. N. and Patricola, C. M.: Diversity of ENSO Events Unified by Convective Threshold Sea Surface Temperature: A Nonlinear ENSO Index, Geophysical Research Letters, 45, 9236–9244, https://doi.org/10.1029/2018GL079203, http://doi.wiley.com/10.1029/2018GL079203, 2018.

Zarzycki, C. M. and Ullrich, P. A.: Assessing sensitivities in algorithmic detection of tropical cyclones in climate data, Geophysical Research Letters, 44, 1141–1149, https://doi.org/10.1002/2016GL071606, http://doi.wiley.com/10.1002/2016GL071606, 2017.

Zarzycki, C. M., Thatcher, D. R., and Jablonowski, C.: Objective tropical cyclone extratropical transition detection in high-resolution reanalysis and climate model data, Journal of Advances in Modeling Earth Systems, 9, 130–148, https://doi.org/10.1002/2016MS000775, http://doi.wiley.com/10.1002/2016MS000775, 2017.

Zhu, Y. and Newell, R. E.: A Proposed Algorithm for Moisture Fluxes from Atmospheric Rivers, Monthly Weather Review, 126, 725–735, https://doi.org/10.1175/1520-0493(1998)126<0725:APAFMF>2.0.CO;2, http://journals.ametsoc.org/doi/abs/10.1175/1520-0493{%}281998{%}29126{%}3C0725{%}3AAPAFMF{%}3E2.0.CO{%}3B2, 1998.